CPPC 2021

First Canadian Peptide and Protein Community Virtual Symposium

27–28 MAY 2021 | ONLINE

# Bio-incorporation of TePhe, a tellurium-containing phenylalanine analogue, preserves protein structure & stability

**Yong Jia (Jamie) Bu[1], Mark Nitz[1*]**

[1] Department of Chemistry, University of Toronto, Toronto, ON, Canada M5S 3H6

* Corresponding author: mark.nitz@utoronto.ca

# Abstract:

The heavy chalcogen tellurium ($^{52}$Te) is a versatile element with applications in mass cytometry, fluorescence imaging, structural biology, and more. L-tellurienylalanine (TePhe), a mimic of L-phenylalanine (Phe) in which the phenyl side chain is replaced by a tellurophene ring, can be covalently incorporated into the proteome of prokaryotes and eukaryotes by endogenous translation machinery. We seek to generate proteins with high levels of Phe→TePhe substitutions, verify preservation of protein structure, and ultimately exploit the incorporated Te as handles for crystallographic phasing, NMR spectroscopy, and bio-orthogonal reactivity.

Here we report conditions for the production of a TePhe-containing protein in a standard *E. coli* expression system. Our target for TePhe incorporation is immunoglobulin-binding Protein G B1 domain (GB1), a 56-residue domain containing 2 Phe residues packed against one another within its hydrophobic core. In Phe-deficient media containing an inhibitor of aromatic amino acid biosynthesis, we obtained a GB1 mixture in which approximately 1 in 2 Phe sites were substituted by TePhe. Fractionation by reverse-phase HPLC yielded a GB1 mixture with 85% TePhe substitution. $^{1}$H-$^{15}$N HSQC and circular dichroism spectroscopy data suggest that TePhe effectively mimics Phe and alters the melting temperature of the protein by less than 5 °C.

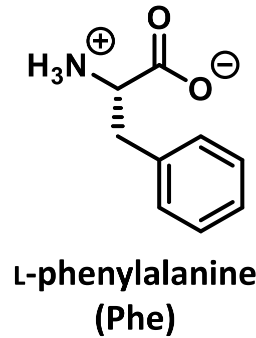


# Keywords

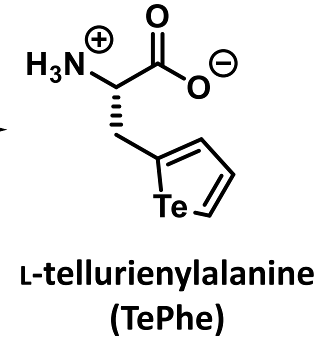
amino acid analogue; organotellurium; metabolic incorporation;
bio-isostere


CPPC
2021

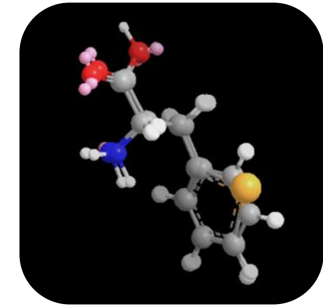
# TePhe is a phenylalanine bio-isostere



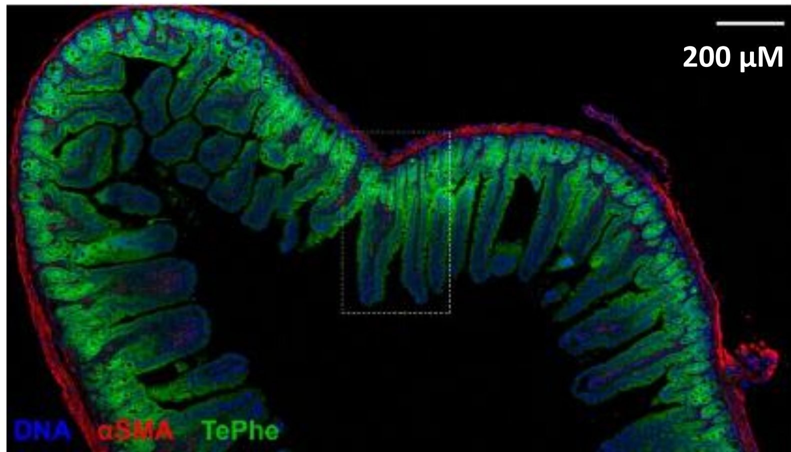

L-phenylalanine
(Phe)

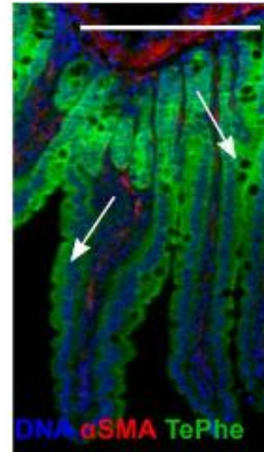



L-tellurienylalanine
(TePhe)



*TePhe as a protein synthesis probe by imaging mass cytometry*




Mass cytometry typically requires <1% Phe→TePhe substitution for robust signal.

***Can we produce proteins with high levels of Phe→TePhe substitutions?***

***How might highly TePhe-substituted proteins differ in structure/function, if at all, from wildtype proteins?***

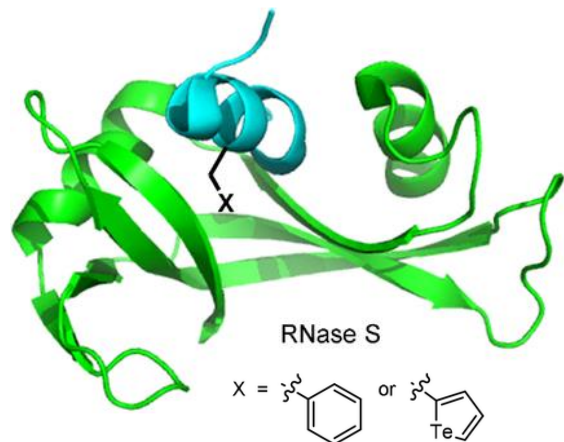
Potential applications in:

- **X-ray crystallography?**
- **NMR spectroscopy?**
- **Novel protein activities?**

CPPC 2021

Bassan, J. *et al. PNAS.* **2019,** *116,* 8155-8160.

# Previous work: semi-synthetic incorporation of TePhe into RNAse S




RNAse S: non-covalent complex between S-peptide (N-terminal 20 residues of RNAse A) and S-protein (C-terminal 104 residues)

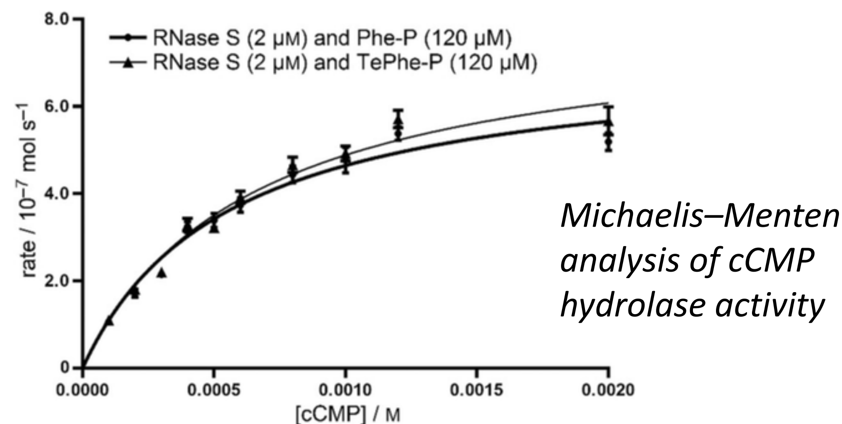
**Phe8 is important for S-peptide association with S-protein!**

*S-peptide* N-**KETAAAKFERQHMDSSTSAA**-C

Phe and TePhe S-peptides were synthesized on solid-phase and complexed with purified S-protein.

**$K_d$ Phe: 1.3 ± 0.5 µM $K_d$ TePhe: 2.63 ± 0.04 µM**

| Mutation at Position 8 | ΔΔG° of association (kJ mol$^{-1}$) |
|---|---|
| Phe → TePhe[1] | 1.8 ± 0.1 |
| Phe → Nle[2] | 9.6 ± 0.4 |
| Phe → Nal[3] | 3.5 |
| Phe → Trp[3] | 7.7 |
| Phe → Tyr[3] | 12.5 |
| Phe → Met[4] | 15 |




*Michaelis–Menten analysis of cCMP hydrolase activity*

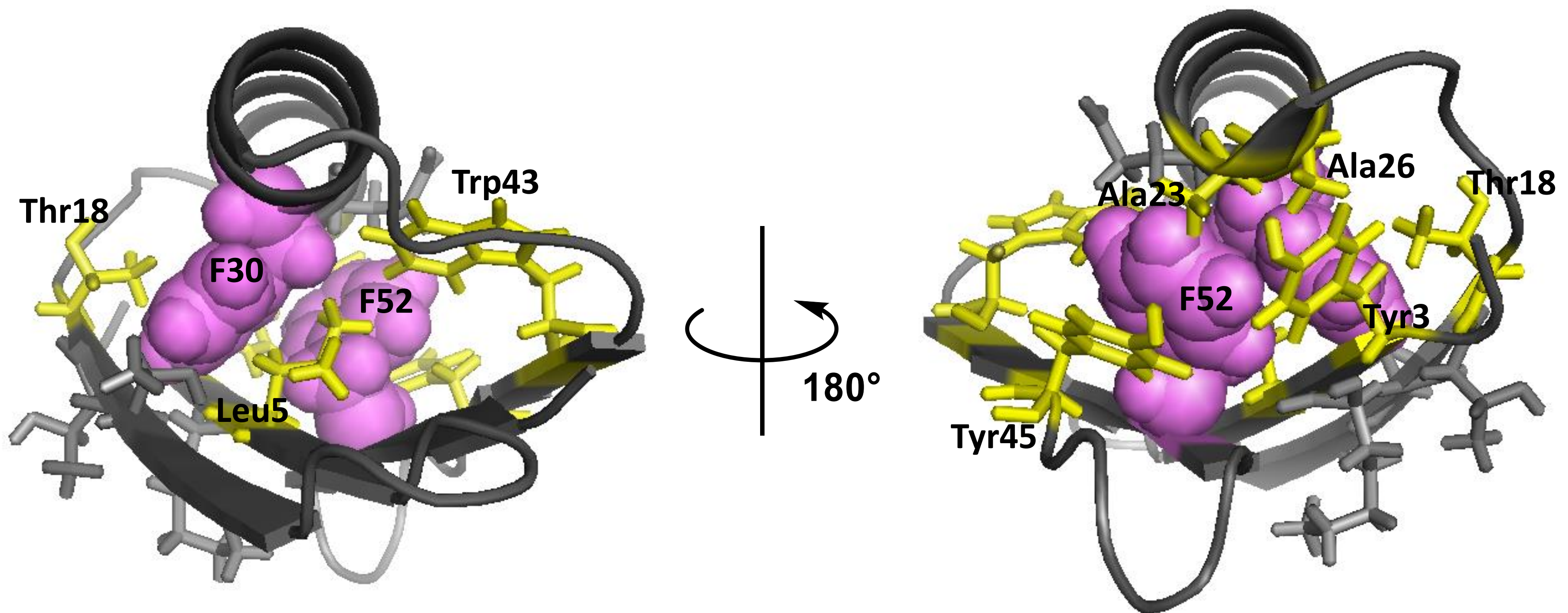
| Parameter | Phe | TePhe |
|---|---|---|
| $k_{cat}$ (s$^{-1}$) | 0.40 ± 0.03 | 0.36 ± 0.03 |
| $K_m$ (mM) | 0.6 ± 0.1 | 0.5 ± 0.1 |
| $K_{cat}/K_m$ (M$^{-1}$ s$^{-1}$) | 620 ± 110 | 650 ± 120 |

1. Vurgun, N. *et al. ChemBioChem.* **2020,** *21,* 1136-1139.
2. Das, M. et al. *Biochemistry* **2005,** *44,* 5923–5930.
3. Bachmann, A. et al. *Proc. Natl. Acad. Sci.* **2011**, *108,* 3952–3957.
4. Ratnaparkhi, G. S. et al. *Biochemistry* **2000,** *39,* 12365–12374.

# Target for TePhe bio-incorporation: GB1



- Domain within Protein G produced by *Streptococcus sp.;* binds heavy chain of immunoglobulin G $F_c$ region
- 56-residue domain containing 2 phenylalanines (F30, F52) which pack against one another in the domain's hydrophobic core




- Very stable, soluble; expresses rapidly and with high yields



Wilton, D. J. *et al. Proteins.* **2008,** *71,* 1432-1440. PDB ID: 2J52.

# Expression of TePhe-containing GB1 in *E. coli*

**BL21 (DE3) *E. coli* with plasmid encoding N-terminally his-tagged GB1 under T7 control (gift from Prof. Joelle Pelletier, UdeM)**

Grow cells in rich media at 37°C to $OD_{600}$ ~0.6

↓

Wash cells with PBS; transfer to M9 minimal media containing 19 canonical AAs + TePhe + glyphosate

↓

Incubate 30 min in expression media to allow TePhe uptake and inhibition of Phe biosynthesis

↓

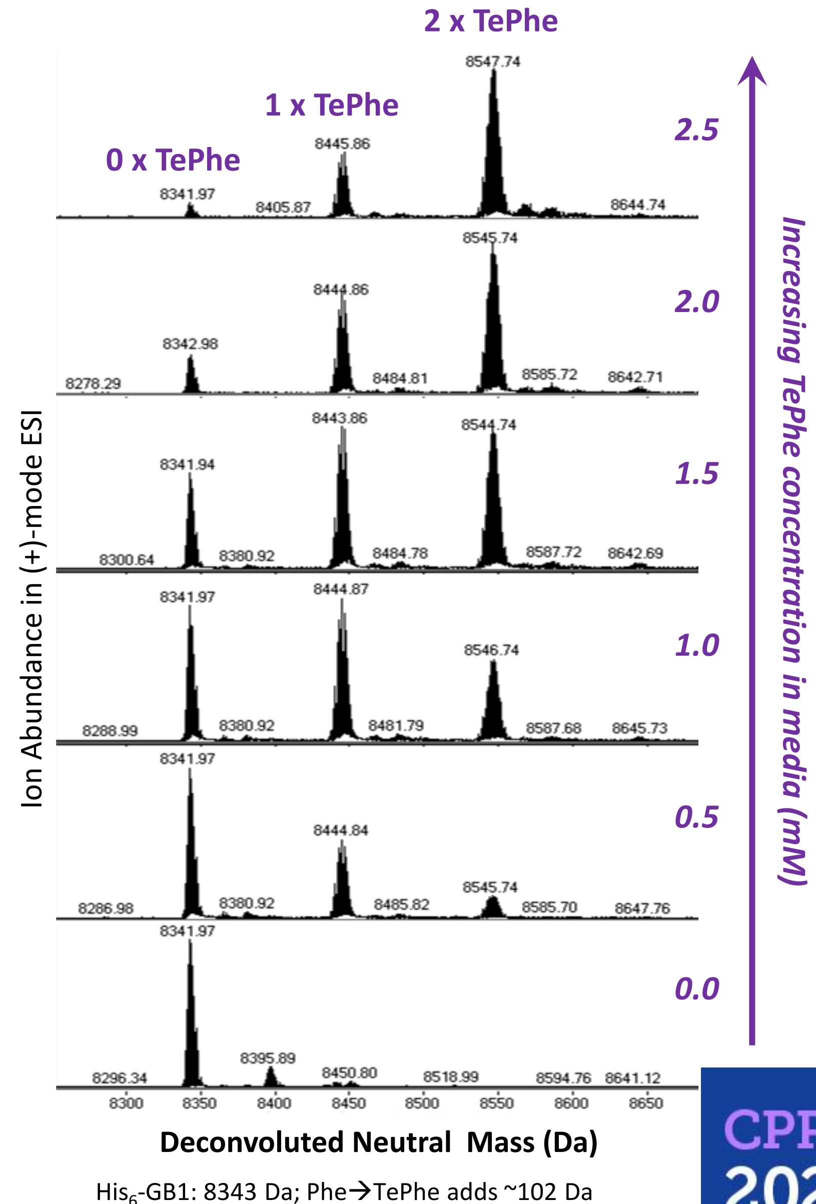
Induce expression with IPTG **in capped tubes (to minimize aeration)**, for 5 hr at 20°C

↓

Lyse cells, perform affinity chromatography

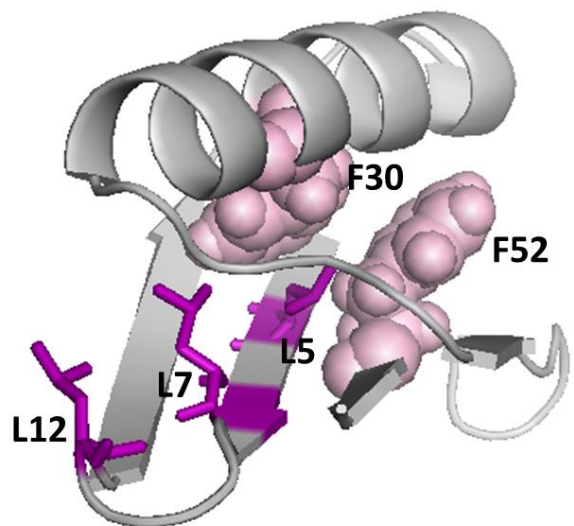
***Amino acid analysis revealed 45% TePhe substitution in GB1 expressed at 2.5 mM TePhe.***

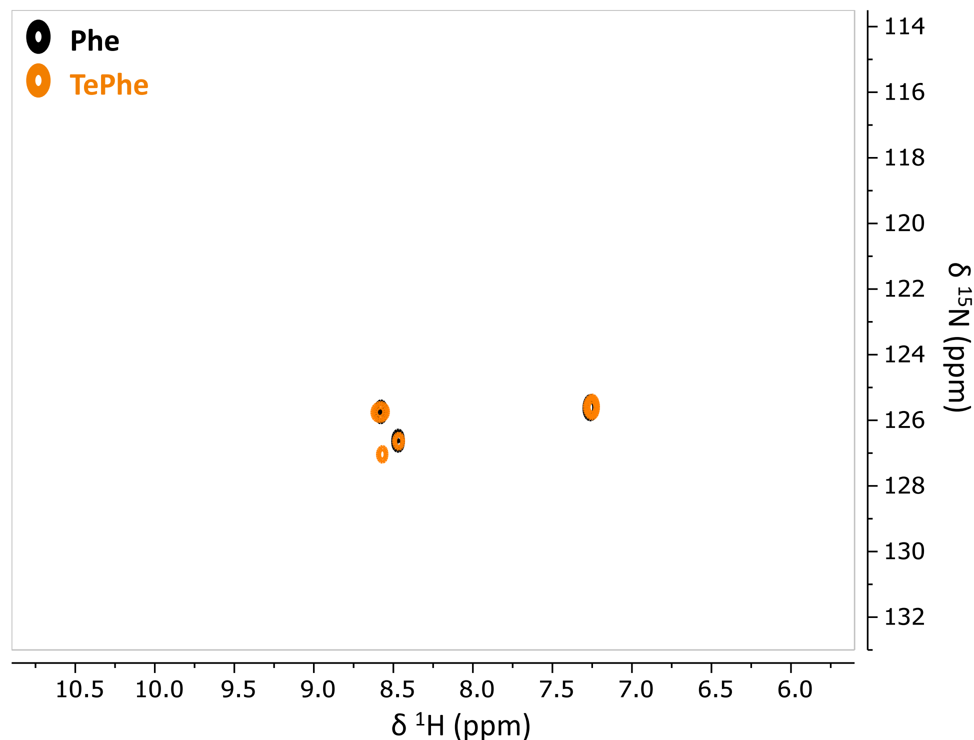
ESI mass spectra of TePhe-substituted $His_6$-GB1



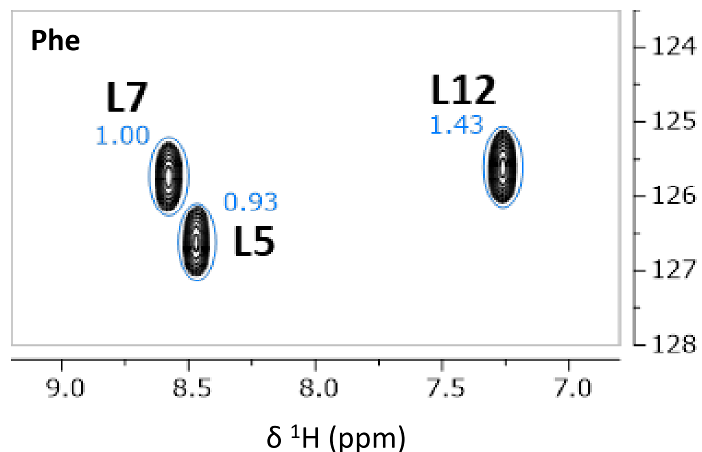

$His_6$-GB1: 8343 Da; Phe→TePhe adds ~102 Da

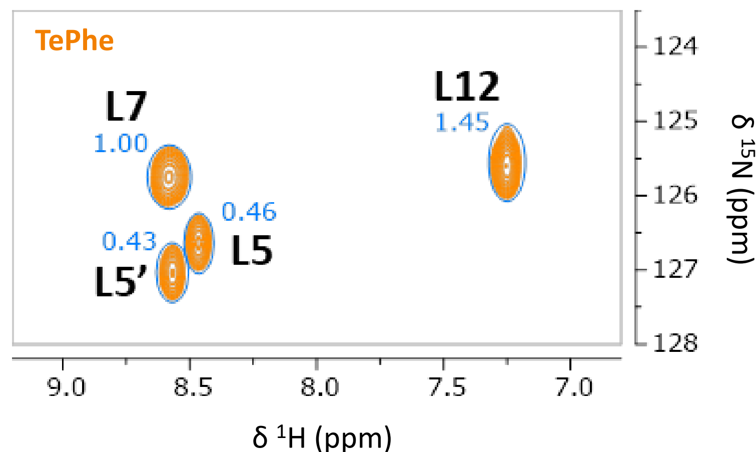
# GB1 leucines as reporters of TePhe substitution effects




*Overlay of $^1$H-$^{15}$N HSQC spectra at 25°C showing $^{15}$N-Leu amides in Phe and TePhe GB1 (10% $D_2O$ in 20 mM phosphate buffer, pH 7.5. $^1$H: 700 MHz, $^{15}$N: 71 MHz.)*




CPPC
2021

# Variable-temperature HSQCs of Phe and TePhe GB1

Temperature (°C): 25 35 45 55 65 75

Phe

L7 L12 L5

124 125 126 127 128 129

9.0 8.5 8.0 7.5 7.0

δ ¹⁵N (ppm)

TePhe

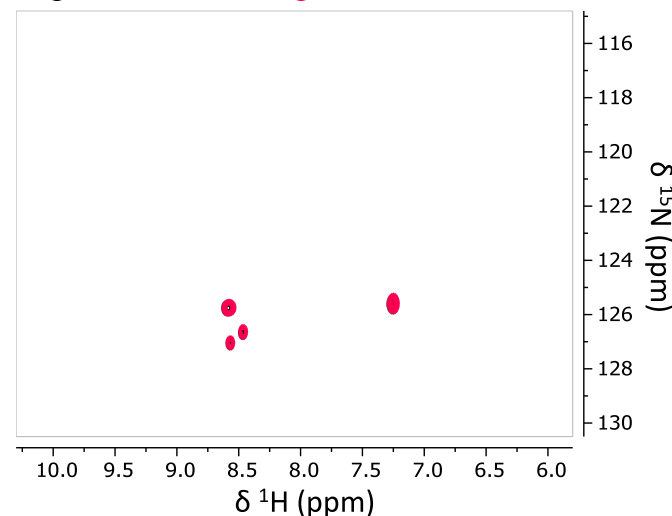
L7 L12 L5 L5'

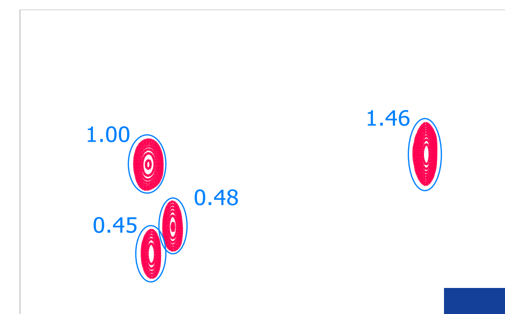
124 125 126 127 128 129

9.0 8.5 8.0 7.5 7.0

δ ¹H (ppm)

*GB1 unfolding is reversible.*
*TePhe GB1 $^{15}$N-Leu amide peaks are recovered upon cooling.*

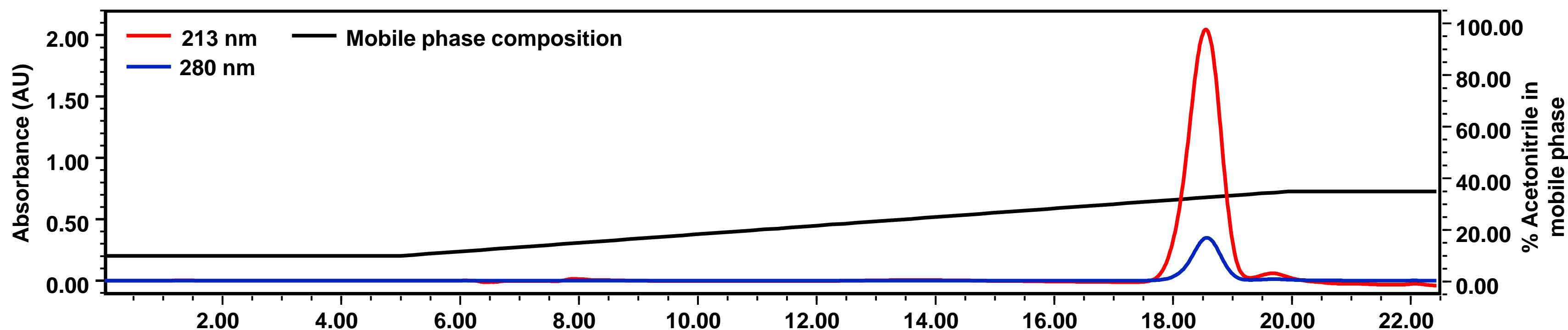


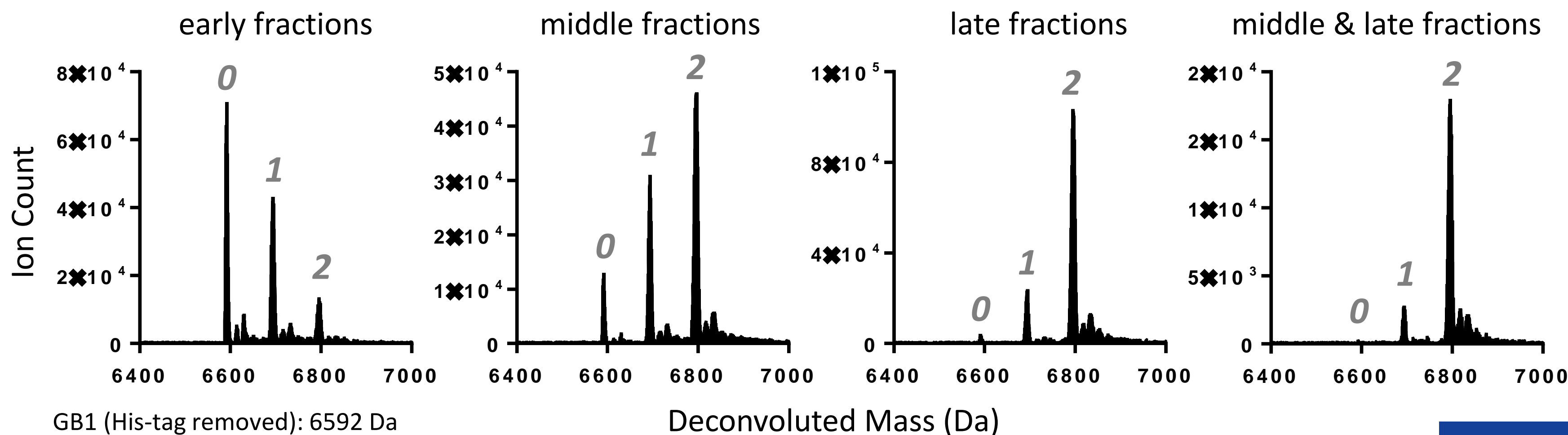

*Variable temperature $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-Leu amides in Phe and TePhe GB1 (10% $D_2O$ in 20 mM phosphate buffer, pH 7.5. $^{1}$H: 700 MHz, $^{15}$N: 71 MHz.)*

# Fractionation of TePhe GB1 by reverse phase HPLC

*Enrichment of TePhe GB1 can be achieved by RP-HPLC as TePhe appears to impart slightly greater affinity for C18 stationary phase.*



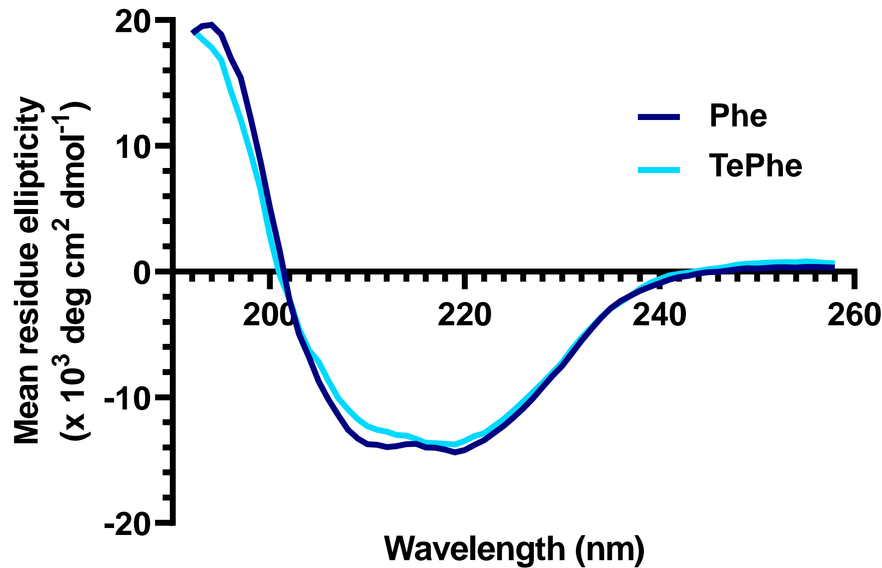

*Prep C18 column, gradient of 10% to 35% acetonitrile in 10 mM tris pH 8; 1 mL/min flow rate.*




**Amino acid analysis of final TePhe-enriched GB1 revealed 85% substitution.**
*(2% non-substituted, 26% 1 x TePhe, 72% 2 x TePhe assuming a random distribution)*

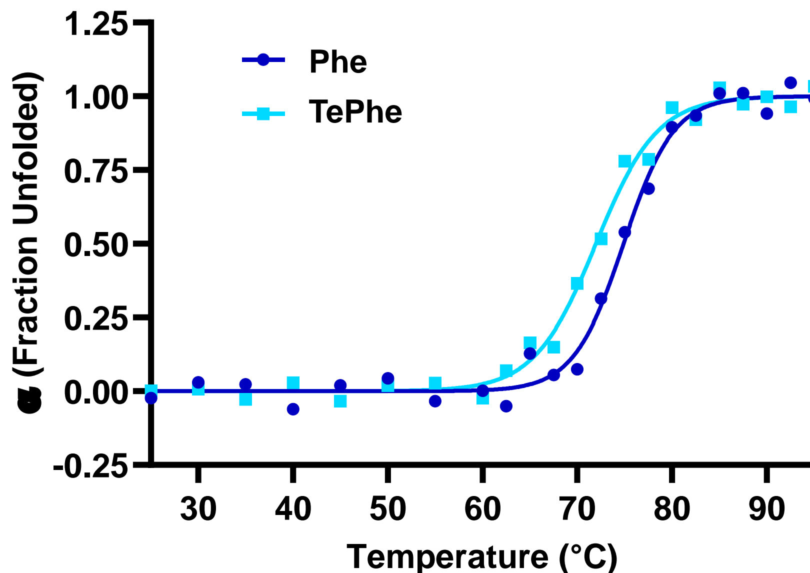
CPPC 2021

# Monitoring thermal denaturation of TePhe GB1 by CD spectroscopy




*Samples prepared in 5 mM phosphate buffer, pH 7.5.*
*Phe: unsubstituted. TePhe: 85% TePhe substituted (HPLC-enriched).*

**CD spectroscopy results suggest that the overall structure and thermal stability of GB1 are largely preserved upon single or double TePhe substitution.**




**Phe $T_m$ = 75.0 ± 1.7 °C**
**TePhe $T_m$ = 72.0 ± 1.7 °C**

*Thermal denaturation monitored by change in mean residue ellipticity at 218 nm (each datapoint is the average of 6 measurements).*

Calfitter software: Mazurenko, S., et al. *Nucleic Acids Res.* **2018,** *46* (W1), W344-W349.

# Conclusions

- Conditions were found enabling approximately 1 in 2 Phe sites in an overexpressed protein to be replaced by TePhe using a standard *E. coli* expression system
- Phe→TePhe substitution is minimally perturbing to the solution-state structure of GB1 as evidence by $^{1}H$-$^{15}N$ HSQC experiments
- GB1 with 85% Phe→TePhe substitution exhibits a very similar far UV CD spectrum to and has a melting point within 5 °C of the wildtype protein

# Future directions

- Bio-incorporate TePhe into larger protein targets with more Phe sites
- Increase level of TePhe bio-incorporation through use of alternate protein expression systems
- Assess utility of TePhe-containing proteins for applications such as crystallography, NMR spectroscopy…

# Acknowledgments

Nitz Group, UofT

- Prof. Mark Nitz
- Dr. Lisa Willis
- Alex Eddenden
- Zach Morrison
- Nicole Potter
- Ehsen Tayyabi
- Dr. Nesrin Vurgun
- All past and present members

Committee Members

- Prof. Andrew Woolley
- Prof. Voula Kanelis
- Prof. Deborah Zamble

Facilities and Services

- UofT AIMS Facility
- UofT CSICOMP-NMR Facility
- SPARC Biocentre (SickKids)

Collaborators and Friends

- Pelletier Lab (University of Montreal)
- Fraser Lab (UCSF)

CPPC 2021